Beyond Routing via Network Coding: An Overview of Fundamental Information-Theoretic Results

Marco Di Renzo, Michela Iezzi, Fabio Graziosi

▶ **To cite this version:**

Marco Di Renzo, Michela Iezzi, Fabio Graziosi. Beyond Routing via Network Coding: An Overview of Fundamental Information-Theoretic Results. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Sep 2010, Turkey. pp. 1-6. hal-00547054

HAL Id: hal-00547054
https://hal.science/hal-00547054

Submitted on 15 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Beyond Routing via Network Coding: An Overview of Fundamental Information–Theoretic Results

(Invited Paper)

Marco Di Renzo
L2S, UMR 8506 CNRS – SUPELEC – Univ Paris–Sud
Laboratory of Signals and Systems (L2S)
French National Center for Scientific Research (CNRS)
École Supérieure d'Électricité (SUPÉLEC)
3 rue Joliot–Curie, 91192 Gif–sur–Yvette (Paris), France
E–Mail: marco.direnzo@lss.supelec.fr

Michela Iezzi and Fabio Graziosi
University of L'Aquila
College of Engineering
Dept. of Electrical and Information Engineering
Center of Excellence DEWS
Via Campo di Pile, 67100 L'Aquila, Italy
E–Mail: {michela.iezzi, fabio.graziosi}@univaq.it

***Abstract*—Since the pioneering research work of Ahlswede *et al.* in 2000, Network Coding (NC) has rapidly emerged as a major research area in electrical engineering and computer science due to its wide applicability to communication through real networks. The many contributions available in the literature to date, ranging from pure theoretical studies on fundamental limits to practical experimentations in real–world environments, offer a clear evidence that the shift in paradigm envisaged by NC might revolutionize the way we manage, operate, and understand the organization of networks. NC allows intermediate nodes of communication networks to combine the information received from multiple links for subsequent transmissions, and offer a powerful and efficient generalization to network information delivery via routing, where network nodes simply store and forward data, and processing is only accomplished at the end nodes. In this paper, we have a twofold objective: i) first, we summarize fundamental information–theoretic results, which, since their publication, have been representing the foundation for all subsequent research in this field, and ii) then, we introduce and summarize the latest results related to the analysis, design, and optimization of the so–called *network error correction coding* problem, which is instrumental for the effective use of NC over lossy, *e.g.*, wireless, networks.**

I. Introduction

Wireless networked systems arise in various communication contexts, and are becoming a bigger and integral part of our everyday life. In today practical networked systems information delivery is accomplished through *routing*: network nodes simply store and forward data, and processing is accomplished only at the end nodes. Network Coding (NC) is a recent field in electrical engineering and computer science that breaks with this assumption: instead of simply forwarding data, intermediate network nodes may recombine several input packets into one or several output packets. NC offers the promise of improved performance over conventional network routing techniques. In particular, NC principles can significantly impact the next–generation wireless *ad hoc* and sensor networks, in terms of both energy efficiency and throughput [1]–[3].

Among the many advantages offered by NC, notable examples are as follows:

- By allowing intermediate nodes in a network to combine information streams and extract the information at the receivers, the throughput of multicast connections can be increased. The "butterfly network" is the most famous example demonstrating this benefit [4].
- In a wireless environment, NC can be used to offer benefits in terms of battery life, wireless bandwidth, and delay. A simple example showing these benefits is the so–called two–way relay channel [5].
- Sending linear combinations of packets instead of uncoded data offers a natural way to take advantage of multipath diversity for security against wiretapping attacks [2].

However, the theory of NC is still in its infancy, and little is known about practical algorithms for exploiting coding in real networks. For example, while most of the existing NC theory assumes error–free links, in practice these links are usually prone to errors arising from, *e.g.*, noise, fading, and interference. Moreover, employing NC introduces some challenges, like, *e.g.*: i) complexity, since it requires the nodes in the network to have additional functionalities, and ii) security concerns, since a sound deployment of NC would require to put in place mechanisms that allow NC operations without affecting the authenticity of the transmitted data. Furthermore, when applying NC to a wireless context we need to take into account that the wireless medium is highly unpredictable and inhospitable to applying the existing NC algorithms, which have been mostly designed by assuming wired networks as the blueprint.

Motivated by the appealing potentials of NC over routing, the aim of this paper is two–fold: i) first, we offer a careful overview of key fundamental results for NC, which represent the starting point for all research activities currently being conducted in this field, and ii) second, we provide a detailed and up–to–date description of the problem of designing codes that combine error–correction coding with NC. In particular, this latter research topic extends significantly beyond the realm of classical coding theory, is extremely challenging, requires new methods and new ideas, and is receiving the interest of the research community only very recently.

The reminder of this paper is organized as follows. In Section II, fundamental information–theoretic results for NC are briefly but carefully outlined. In Section III, we introduce the concept of error–correction codes in projective spaces, and describe its importance for the robust application of NC over lossy networks. Finally, Section IV concludes the paper.

II. Fundamental Information–Theoretic Results

The main NC theorem was introduced for multicast connections in 2000 [4]. In particular, in [4] it has been shown that the capacity of multicast networks (*i.e.*, the maximum

number of packets that can be sent from the source to a set of terminals per time unit) can be achieved by coding within the network, *i.e.*, by allowing intermediate network nodes to not only store and forward data, but also combine the incoming information flows. This result introduces a fundamental shift in paradigm in the way we conceive and design networks. For long time, it has been assumed that when multiple receivers need to simultaneously send data through a network they have to share its resources, thus leading to a reduction of the rate that can be achieved by each of them. On the contrary, with NC each receiver can achieve the same rate as it is the sole node having access to the resources of the network.

A. The Main NC Theorem: Ahlswede et al. [4]

The fundamental result for NC has been proved in [4] and can formally be stated as follows:

Theorem 1: Let us consider a network represented by a directed acyclic graph $G = (V, E)$, with V denoting the set of vertices and $E \subset V \times V$ being the set of edges. It is assumed that each edge has unit capacity and that there are r unit–rate sources on the same vertex of the graph that wish to deliver data to d receivers. Let us also assume that the value of the min–cut to each receiver is r. Then, there exists a multicast transmission scheme over a finite field $\mathbb{F}_q$ in which the intermediate nodes of the network combine their incoming information symbols over $\mathbb{F}_q$, so that the information emitted by the sources can be simultaneously delivered to the d receivers at a rate equal to r. □

In particular, this theorem can be regarded as the Max–Flow Min–Cut theorem for network information flow [6], [7], which is summarized in the following definition and theorem:

Definition 1: Let us consider a network represented by a graph $G = (V, E)$. Let V and $E \subset V \times V$ denote the set of vertices and the set of unit–capacity edges, respectively. Also, let $S \in V$ be a source node of the network that wishes to transmit data to a destination node $D \in V$. Then, the following definitions hold:

- A *cut* between S and D is a set of graph edges whose removal disconnects S from D;
- A *min–cut* is a cut with the minimum value;
- The *value* of a cut is the sum of the capacities of the edges in the cut. □

Theorem 2: Let us consider a network represented by the graph under the assumptions of the definition above. If the min–cut between S and D is equal to r, then the information from S to D can be sent at a maximum rate of r. □

B. From NC to Linear NC: Li et al. [8]

The theorems above clearly state that by allowing the mixing of data at the intermediate nodes of the network the information from a source node to a destination node can be delivered through the network at a maximum rate equal to the min–cut between them. The landmark result in [4] does not put any constraints on how to combine the incoming packets at each intermediate node for achieving the max–flow. An important result about this latter issue was achieved by Li *et al.* in 2003 [8]. As a matter of fact, the authors have shown that, for multicast networks, *linear coding* at the intermediate nodes suffices to achieve the capacity limit, which is the max–flow bound from the source to each receiving node. In [8], the concept of *linear–code multicast* has been introduced and explicit code constructions for multicast networks have been provided for both acyclic and cyclic networks.

C. Using Algebra for the Design of Linear NC: Koetter and Médard [9]

An important problem for linear NC is to construct the network code or, in other words, to find the coefficients of the encoding functions at the intermediate nodes so that each receiver of the multicast network problem can retrieve the source messages from the received packets. The sets of coefficients satisfying this decoding condition yields a solution of the NC problem, which is, in turn, called solvable. A general but simple and systematic answer to this important question was given by Koetter and Médard in 2003 [9], by resorting to a powerful algebraic framework. In particular, the main result in [9] establishes connections between the solutions of NC problems and the solutions of systems of linear equations. In what follows, we summarize the main results in [9], which are currently being used extensively by many researchers in the field. Let us emphasize here that the design method in [9] is useful for networks with and without delays (or cycles). However, for the sake of simplicity, we analyze here only networks with no delays.

Before introducing the main theorems, let us start with some notations and definitions needed to describe in a compact fashion the NC problem under analysis. The network is represented by a directed graph $G = (V, E)$, with V denoting the set of network nodes (vertices) and $E \subset V \times V$ being the set of network links (edges). We assume the information is being sent noiselessly from node i to node j for all $(i, j) \in E$.

Definition 2: Given a generic link $(i, j) \in E$, node i and node j are called origin and destination, respectively. For any generic link $l \in E$, the origin and destination of l are denoted by $o(l)$ and $d(l)$, respectively. Via NC the information transmitted on a link $l \in E$ is obtained as a coding function of the packets previously received at $o(l)$. □

Definition 3: Let r and d be the number of sources processes being transmitted through the network and the number of receivers, respectively. Also, let source and destination nodes be denoted by $\{a(1), a(2), \ldots, a(r)\}$ and $\{\beta_1, \beta_2, \ldots, \beta_d\}$, respectively. Furthermore, let the source information processes, the destination output processes, and the information processes transmitted in each link be sequences of length–u vectors of bits that are treated as elements of the finite field $\mathbb{F}_q$ with $q = 2^u$. Then, using NC the i–th output process, which is denoted by $Z_{\beta,i}$, at the generic destination node β is a linear combination of the information processes on its terminal link, as follows:

$$Z_{\beta,i} = \sum_{\{j:d(j)=\beta\}} b_{\beta,i,j} Y_j \tag{1}$$

where Y_j is the information process transmitted on link j, which is obtained as a linear combination in $\mathbb{F}_q$ of: 1) the inputs of link j, *i.e.*, the source processes X_i for which $a(i) = o(j)$, and 2) the random processes Y_l for which $d(l) = o(j)$. Accordingly, we have:

$$Y_j = \sum_{\{i:a(i)=o(j)\}} a_{i,j} X_i + \sum_{\{l:d(l)=o(j)\}} f_{l,j} Y_l \tag{2}$$

The set of coefficients $\{a_{i,j}, f_{l,j}, b_{\beta,i,j}\}$ are suitably chosen numbers in $\mathbb{F}_q$ that define the NC problem: they are a solution of the NC problem when all desired connections can be successfully established by the network. □

From the two definitions above, the following two important theorems have been introduced in [9]:

Theorem 3: Let the coefficients $\{a_{i,j}\}$ be collected into a[1] $r \times |E|$ matrix $\mathbf{A}$, the coefficients $\{b_{\beta,i,j}\}$ be collected into a $r \times |E|$ matrix $\mathbf{B}_\beta$, and the coefficients $\{f_{l,j}\}$ be collected into a $|E| \times |E|$ matrix $\mathbf{F}$. The tuple $(\mathbf{A}, \mathbf{F}, \mathbf{B}_{\beta_1}, \mathbf{B}_{\beta_2}, \ldots, \mathbf{B}_{\beta_d})$ is called *linear network code*.

Then, the mapping from the source processes[2] $\mathbf{X} = [X_1, X_2, \ldots, X_r]^T$ to the output processes $\mathbf{Z}_\beta = [Z_{\beta,1}, Z_{\beta,2}, \ldots, Z_{\beta,r}]^T$ at a generic receiver[3] β is as follows:

$$\mathbf{Z}_\beta = \left(\mathbf{A}\mathbf{G}\mathbf{B}_\beta^T\right)\mathbf{X} \tag{3}$$

where $\mathbf{G} = (\mathbf{I} - \mathbf{F})^{-1}$ and $\mathbf{I}$ is the identity matrix. □

Theorem 4: Given a multicast connection problem, the linear network code $(\mathbf{A}, \mathbf{F}, \mathbf{B}_{\beta_1}, \mathbf{B}_{\beta_2}, \ldots, \mathbf{B}_{\beta_d})$ over the field $\mathbb{F}_q$ is said to be a solution of it if the transfer matrix $\mathbf{M}_{\beta_k} = \mathbf{A}\mathbf{G}\mathbf{B}_{\beta_k}^T$ has full rank r for each receiver β_k with $k = 1, 2, \ldots, d$. In such a case, the multicast connection problem is called feasible. □

An important aspect behind the construction of linear network codes for NC problems is to understand the size of the Galois field $\mathbb{F}_q$, *i.e.*, q, to get a feasible multicast connection problem. This is a crucial aspect from a practical point of view since the larger the Galois field is, the more computational complex the operations performed by the network are. As a matter of fact, for a given q, the algebraic operations of NC are performed on codewords of length $u = \log_2(q)$. By exploiting the algebraic approach in [9], the following important lower bound holds [10]:

Theorem 5: For a feasible multicast connection problem with independent or linearly correlated sources and d receivers, in both the acyclic delay–free case and the general case with delays, there exists a linear network code $(\mathbf{A}, \mathbf{F}, \mathbf{B}_{\beta_1}, \mathbf{B}_{\beta_2}, \ldots, \mathbf{B}_{\beta_d})$ in a Galois field $\mathbb{F}_q$ if $q > d$. □

D. Random Linear NC: Ho et al. [10]

The fundamental theorems derived in [9] are driven by the main goal of computing a *deterministic* linear network code that solves the multicast connection problem. In other words, the coefficients $\{a_{i,j}, f_{l,j}, b_{\beta,i,j}\}$ are chosen so that, over noiselessly networks, the source processes can deterministically be retrieved at all destination nodes with unit probability. However, this requirement implicitly assumes the adoption of network codes that are either planned or are known by a given central authority, which reduces to having a centralized network architecture. A sound answer to this problem has been given in [10], where the authors have proposed a distributed linear NC approach for a general multi–source multicast network. The key element of the method in [10] is to let the network nodes randomly select the encoding functions, *i.e.*, the coefficients of the linear combinations, at the intermediate nodes. In the light of its operating principle, this method has been called *random linear NC*. Unlike deterministic linear NC, random linear NC cannot achieve the multicast capacity with unit probability, but the multicast connection problem turns out to be feasible with probability exponentially approaching one with the code length $u = \log_2(q)$. The main theorem in [10] can be stated as follows:

Theorem 6: Let us consider a multicast connection problem with independent or linearly correlated sources, d destination nodes, and a linear network code in which some or all network code coefficients $\{a_{i,j}, f_{l,j}, b_{\beta,i,j}\}$ are chosen uniformly at random over a finite field $\mathbb{F}_q$ with $q > d$, and the remaining code coefficients, if any, are fixed. If there exists a solution to the network connection problem with the same values of the fixed code coefficients, then the probability that the random network code is valid for the problem is at least $(1 - d/q)^\eta$, where η is the number of links with associated random coefficients. □

This important theorem clearly shows that: i) the larger the set of links with random coefficients is (*i.e.*, η), the smaller the probability that the random network code is valid is, and ii) the larger the size of the Galois field is (*i.e.*, q and so the length of the network code $u = \log_2(q)$), the bigger the probability that the random network code is valid is.

E. Distributed NC: Chou et al. [11]

On the practical side, an important result for NC has been achieved by *Chou et al.* [11]. The theoretical bounds predicted in [4], [8], and [9] rely on the assumption of some centralized knowledge of both the network topology and the coding functions at the intermediate nodes, which are required in order to design the network code and to decode the information at the destination. However, in real networks such a centralized knowledge might be very difficult to be achieved in practice. As mentioned above, a first step towards the design of distributed NC has been made in [10] with the introduction of random NC. However, the randomly chosen coding functions in [10] need to be delivered to the destination nodes to allow them to retrieve the received network–coded source processes. This is a fundamental issue to be accounted for to design a totally distributed network code. Furthermore, in practical networks the information is likely to travel asynchronously and be subject to random delays, which should be taken into account during the network–encoding process.

These fundamental and practical problems have been addressed in [11], and a clever distributed coding scheme has been proposed, which obviates the need of the centralized knowledge of the graph topology, the encoding and decoding functions at each intermediate node, and does not rely on any assumptions of synchronous transmission of the information through the network. The fundamental ideas behind the distributed NC solution in [11] are as follows:

- The centralized knowledge of the network topology and the encoding/decoding functions can be avoided by using a *data–aided*–like transmission scheme, which consists in including within each outgoing data packet flowing on an edge of the network a packet header the describes the coefficients of the linear combination of the incoming packets it contains. The coefficients of the linear

[1] $|\cdot|$ denotes the cardinality of a set.

[2] $(\cdot)^T$ denotes the transpose operator.

[3] Let us emphasize here that β can take values in the set $\beta \in \{\beta_1, \beta_2, \ldots, \beta_d\}$.

combination are known as *global encoding vectors*. This way, these latter vectors, which are needed to decode the data received at any receiver, can be found in the arriving packets themselves. With the cost of a reasonable overhead, this approach can offer a totally decentralized solution to NC over networks. In practice, this scheme can be simply accomplished by pre–pending the canonical basis vector to each source vector and processing the resulting packet at each node as foreseen by the NC paradigm.

- The problem of synchronizing the packets related to the same set of source vectors to correctly perform both encoding and decoding can be addressed by introducing the concept of *generation*. In practice, all packets related to the same set of source vectors are said to belong to the same generation, and are tagged with the same generation number that is pre–pended in the packet header. The proposal in [11] allows the combination of packets belonging to only the same generation and adopts a buffering mechanism to accomplish the synchronization of incoming and outgoing packets.

III. Information–Theoretic Results for Error Control/Correction Coding

Besides the many potential advantages and applications of NC over classical routing (see, *e.g.*, [2], [3]), the NC principle is not without its drawbacks. A fundamental problem that NC needs to face over lossy networks is the so–called *error control problem*: corrupted packets injected by some intermediate nodes might propagate through the network until the destination, and might render impossible to decode the original information. As a consequence, it is instrumental to develop efficient techniques to counteract this error propagation effect, while retaining the main benefits of NC over routing. In particular, in this section special emphasis will be given on the so–called *subspace coding* or *coding in projective spaces* method [12], which is a new and recently proposed approach that promises to be very helpful in this situation.

A. NC vs. Routing: Motivation of the Error Control Problem

In contrast to routing, the error control problem is crucial in NC due to the algebraic operations performed by the internal nodes of the network. As a matter of fact, the mixing of packets within the network makes every packet flowing through it statistically dependent on other packets: even a single erroneous packet might affect the correct detection of all other packets. On the contrary, the same error in networks using just routing would affect only a single source–to–destination path. Broadly speaking, possible errors in NC might arise for three main reasons [13]: i) *erasures*, which lead to insufficiently received packets at the destination to solve the NC problem and retrieve the transmitted messages, ii) *errors*, which are due to using, for complexity and practical reasons, not powerful enough link–to–link error–correction codes or are caused by the need to avoid a retransmission of all corrupted packets, and iii) the presence of intentional *jammers*, who might introduce erroneous packets at the application layer, which might be difficult to be recovered by the destination node. In such a context, the conventional approach to drop all erroneous packets detected at the physical layer might be very sub–optimal for several reasons, *e.g.*, i) this may lead to insufficiently received packets for decoding and may be very spectrally inefficient, ii) even packets with errors could be a source of redundancy that may help the decoding process at the destination node, and iii) even though some bits are wrong, some parts of the packets could still be error–free and could be exploited via some joint source–channel decoding methods to correct the wrong bits.

B. Towards Network Error Correction

The first landmark approaches to the design of error control codes for network–coded systems have been presented in [14]–[16]. In these papers, the authors have introduced the concept of *network error correction*, whose main idea is to design the network code so that it can be used for error correction. The underlying idea is to exploit the network code for protecting the messages transmitted through the network from distributed errors occurring over the individual links, which are not assumed to be error–free. Network error correction generalizes the usual link–to–link error correction methods adopted in conventional networks.

The concept of network error–correction code in [14] is based on the following two definitions:

Definition 4: Given a network, an error is said to occur if a symbol at the output of a link is different from the corresponding output symbol. A (distributed) τ–error is said to occur if the total number of errors that occur in all channels of the network is equal to τ. □

Definition 5: A network code is t–error correcting if it can correct all τ–errors for $\tau \leq t$, *i.e.* if the total number of errors in the network is at most t. In such a case, the source message can be recovered by all destination nodes. □

Broadly speaking, the method introduced in [14]–[16] considers the design of a network code as part of an error control problem. Moving from the original idea of network error correction introduced in [14], many subsequent papers have investigated this problem with the main aim of computing fundamental performance bounds, and proposing code constructions and efficient decoding algorithms for network error–correction codes. Notable examples along this line are [13], [17]–[19], and references therein.

C. Coding in Projective Spaces: A New Look at Error–Correction Codes for Linear Random NC

A radical shift in paradigm on the design of error–correction codes for random NC has been introduced by Koetter and Kschischang in [12], who conceived the principle of coding for operator channels. This clever idea has originated an active field of research that is also known as subspace coding or, more recently, as error–correction code design in projective spaces (see, *e.g.*, [20]). The main idea behind [12] resides in recognizing that the natural transmission model of random NC consists of inputs and outputs that are subspaces of a given vector space. The interesting feature and main difference of the method introduced in [12], with respect to previous approaches available in the literature, is to be oblivious to both the network topology and the particular network code. In other words, the method introduced in [12] seeks to design an outer code that can be applied end–to–end without requiring any modifications on (or even the knowledge of) the underlying network code.

The basic idea is to encode the information in the choice, at the transmitter, of a vector space (rather than a vector), and to design, at the receiver, a suitable algorithm to reconstruct the subspace sent by the transmitter in the presence of different kinds of errors.

1) Motivation of Coding in a Vector Space: The theoretic motivation behind the design of error–correction codes in projective spaces relies on the algebraic formulation introduced by Koetter and Médard in [9]. Similar to (3), let $\mathbf{X}$ and $\mathbf{Z}_\beta$ denote the matrices containing the messages at the input and at the output of a network performing NC at the intermediate nodes, respectively. By restricting these nodes to perform only linear operations [8], in lossy networks $\mathbf{X}$ and $\mathbf{Z}_\beta$ can be related as follows:

$$\mathbf{Z}_\beta = \mathbf{M}_\beta \mathbf{X} + \mathbf{\Xi} \tag{4}$$

where, similar to (3), $\mathbf{M}_\beta$ is the matrix corresponding to the overall linear transformations applied by the network, and $\mathbf{\Xi}$ is the matrix of the network–coded error packets injected into the network and due to possible errors over individual links.

Since in random NC the matrix $\mathbf{M}_\beta$ of the overall linear transformation applied by the network is unknown [10], it follows that, even in the absence of errors, the only property of the transmitted packets that is kept invariant after propagation through the channel model in (4) is the product $\mathbf{M}_\beta \mathbf{X}$, which is the row space of $\mathbf{X}$. In other words, from the point of view of the destination nodes, any of the possible generating sets for the space $\mathbf{M}_\beta \mathbf{X}$ are equivalent. As a consequence, the conventional link–to–link code design, which foresees the transmission of the information via a suitable design of $\mathbf{X}$, needs to be modified and generalized to convey the information via the vector space spanned by the row space of $\mathbf{X}$. This is the underlying and fundamental motivation behind subspace coding.

2) Encoding, Decoding, and Operator Channel: Mathematically speaking, the encoding and decoding processes resulting from the motivating idea behind subspace coding can be stated as follows: i) the source node selects a subspace $\mathcal{V}$, which belongs to an ambient space $\mathcal{W}$ over a given Galois field, to be transmitted over the channel model in (4), ii) the channel transforms the subspace $\mathcal{V}$ into the subspace $\mathcal{U}$, which still belongs to $\mathcal{W}$ due to the vector space preserving properties of linear NC, and iii) the destination node receives $\mathcal{U}$ from which it tries to infer $\mathcal{V}$. By taking into account (4), the input and output spaces $\mathcal{V}$ and $\mathcal{U}$, respectively, can be related to each other by introducing the concept of *operator channel*:

Definition 6: An operator channel associated to an ambient space $\mathcal{W}$ is a channel whose input, $\mathcal{V}$, and output, $\mathcal{U}$, are subspaces of $\mathcal{W}$ that can be related as follows:

$$\mathcal{U} = \mathcal{H}_k(\mathcal{V}) \oplus \mathcal{E} \tag{5}$$

where $\oplus$ denotes the direct sum, $\dim(\cdot)$ denotes the dimension of a vector space, $\mathcal{E}$ is the error space, $k = \dim(\mathcal{V} \cap \mathcal{U})$, and $\mathcal{H}_k(\cdot)$ is the so–called *erasure operator* that is defined in the following way: $H_k(\mathcal{V}) = \mathcal{V}$ if $\dim(\mathcal{V}) \leq k$, while $H_k(\mathcal{V})$ returns a random k–dimensional subspace of $\mathcal{V}$ if $\dim(\mathcal{V}) > k$.

The operator channel in (5) takes into account that packet erasures might happen when transmitting the information through the network, so that $\mathbf{M}_\beta \mathbf{X}$ in (4) can only generates a subspace of the row space of $\mathbf{X}$ in (4). In transforming $\mathcal{V}$ to $\mathcal{U}$, the operator channel is said to introduce $\rho = \dim(\mathcal{V}) - k$ erasures and $t = \dim(\mathcal{E})$ errors. □

3) Performance Guarantees – The Main Theorem: The fundamental result in [12] is concerned with the definition of a minimum–distance decoder and a suitable distance metric to retrieve $\mathcal{V}$ from $\mathcal{U}$, along with the understanding of the combined error and erasure correction capabilities of subspace coding. Two fundamental results are summarized in the theorems as follows [12]:

Theorem 7: Let us consider the operator channel in (5), the destination node can recover $\mathcal{V}$ from $\mathcal{U}$ by using the minimum distance decoder as follows:

$$\hat{\mathcal{V}} = \arg\min_{\mathcal{V} \in \mathcal{W}} \{d_S(\mathcal{V}, \mathcal{U})\} \tag{6}$$

where $d_S(\cdot,\cdot)$ is the subspace distance defined as:

$$d_S(\mathcal{V}, \mathcal{U}) = \dim(\mathcal{V}) + \dim(\mathcal{U}) - 2\dim(\mathcal{V} \cap \mathcal{U}) \tag{7}$$

□

Theorem 8: Let us consider the operator channel in (5), the minimum distance decoder in (6) guarantees perfect decoding capabilities, *i.e.*, $\hat{\mathcal{V}} = \mathcal{V}$, provided that:

$$2t + 2\rho \leq d_S(\mathcal{W}) \tag{8}$$

where $d_S(\mathcal{W})$ is the minimum subspace distance as follows:

$$d_S(\mathcal{W}) = \min_{\substack{\mathcal{V}_1, \mathcal{V}_2 \in \mathcal{W} \\ \mathcal{V}_1 \neq \mathcal{V}_2}} \{d_S(\mathcal{V}_1, \mathcal{V}_2)\} \tag{9}$$

with $\mathcal{V}_1$ and $\mathcal{V}_2$ being arbitrary subspaces in $\mathcal{W}$. □

D. Recent Advances on Coding in Projective Spaces

Besides introducing the fundamental theory for constructing error–correction codes in projective spaces in [12], today known as "KK codes", Koetter and Kschischang have also introduced a Reed–Solomon–like construction and have described a Sudan–style minimum–distance decoding algorithm for the new family of subspace codes. Furthermore, the class of constant–dimension codes has been introduced and investigated. Soon after [12], several contributions have appeared in the literature with the goal of generalizing and improving the original idea. Relevant results are [21]–[29]. Due to space constraints, a comprehensive treatment of the main theorems of this promising research field is not possible in the present paper. So we limit ourselves to providing an, to the best of our knowledge, up–to–date reference list and a very short summary of the contribution of each paper. Interested readers might find further details by directly referring to the papers.

In [21], the authors study optimal constant–dimension codes for their application to NC, and show that Steiner structures are optimal constant–dimension codes. Two Johnson–type bounds are also computed. In [22], several new codes and bounds for the subspace metric introduced in [12] are derived. In [23], a large class of constant–dimension subspace codes is investigated. It is shown that codes in that class can be easily constructed from rank–metric codes, while preserving their distance properties. Moreover, it is shown that minimum distance decoding of such subspace codes can be reformulated as a generalized decoding problem for rank–metric codes where partial information about the error is available. Furthermore,

for the important family of maximum rank–distance codes known as Gabidulin codes, an efficient decoding algorithm is proposed. In [24], the authors construct many new constant–dimension codes with a larger number of codewords than previously known codes. In [25], the authors study bounds and code constructions for the family of codes in [12] targeting the correction of insertions/deletions. In [26], the authors analyze the geometrical properties of rank–metric codes. They derive upper and lower bounds on the minimum cardinality of a code with a given rank covering radius and show that the proposed geometrical properties and bounds can be significant to the design, decoding, and performance analysis of rank–metric codes. In [20], a novel multilevel coding approach to construct codes in the projective space is presented. The method uses four tools: an appropriate constant–weight code, the reduced row echelon form of a linear subspace, the Ferrers diagram related to this echelon form, and rank–metric codes related to the Ferrers diagram. The authors show that the codes proposed in [12] are a special case of the proposed family of codes. In [27], the error correction problem in both coherent and non–coherent NC is considered under an adversarial model. In particular, as far as non–coherent NC is concerned, the authors introduce a different metric with respect to [12], and prove that it yields a measure of code performance that is more precise, when a non–constant–dimension code is used, than [12]. The new metric is called injection metric. In [28], the authors introduce a Gilbert–Varshamov bound for the codes constructed in [27] according to the definition of injection metric. Moreover, the construction framework in [20] is exploited to obtain new non–constant–dimension codes, which are shown to contain a larger number of codewords with respect to comparable codes designed for the subspace metric. Finally, in [29] the authors address the very important problem of understanding if the codes introduced in [12] are feasible and suitable for hardware implementations. They show that the construction of these codes over small fields and with limited error–correction capabilities is not only feasible, but the resulting codes can achieve a high throughput.

IV. Conclusion

In this paper, we have provided an overview of important information–theoretic results for NC and have highlighted that most of the existing literature on NC assumes that the links in the network are error–free. While this is a reasonable assumption in theory, this is certainly not true for most real–world networks whose resources (*e.g.*, the transmission power) are limited. So, we have provided an up–to–date survey of the so–called subspace coding approach, which offers a powerful technique to develop efficient network/channel coding algorithms that could be quickly deployed in real–world networks.

Acknowledgment

This work is supported, in part, by the research projects "JNCD4CoopNets" (CNRS – GDR 720 ISIS) and "Re.C.O.Te.S.S.C." of PORAbruzzo, and by DIGITEO.

References

[1] T. Ho, R. Koetter, M. Médard, D. R. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting", *IEEE Int. Symposium Inform. Theory*, pp. 442, June/July 2003.

[2] C. Fragouli, E. Soljanin, *Network Coding Fundamentals*, vol. 2/1, 2007.

[3] C. Fragouli, E. Soljanin, *Network Coding Applications*, vol. 2/2, 2007.

[4] R. Ahlswede, N. Cai, S.–Y. R. Li, and R. W. Yeung, "Network information flow", *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.

[5] S. Zhang, S. Liew, and P. Lam, "Hot topic: Physical layer network coding", *ACM Int. Conf. Mobile Computing and Networking*, pp. 358–365, May 2006.

[6] P. Elias, A. Feinstein, and C. E. Shannon, "Note on maximum flow through a network", *IRE Trans. Inform. Theory*, vol. 2, no. 4, pp. 117–119, Dec. 1956.

[7] L. R. Ford Jr. and D. R. Fulkerson, "Maximal flow through a network", *Canadian J. Mathematics*, vol. 8, pp. 399–404, 1956.

[8] S.–Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding", *IEEE Trans. Inform. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[9] R. Koetter and M. Médard, "An algebraic approach to network coding", *IEEE/ACM Trans. Networking*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[10] T. Ho, R. Koetter, M. Médard, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast", *IEEE Trans. Inform. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.

[11] P. A. Chou, Y. Wu, and K. Jain, "Practical network coding", *Allerton Conf. Commun. Control, Computing*, Oct. 2003.

[12] R. Koetter and F. R. Kschischang, "Coding for errors and erasures in random network coding", *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3579–3591, Aug. 2008.

[13] Z. Zhang, "Linear network error correction codes in packet networks", *IEEE Trans. Inform. Theory*, vol. 54, no. 1, pp. 209–218, Jan. 2008.

[14] N. Cai and R. W. Yeung, "Network coding and error correction", *IEEE Inform. Theory Workshop*, pp. 119–122, Oct. 2002.

[15] R. W. Yeung and N. Cai, "Network error correction, part I: Basic concepts and upper bounds", *Commun. Information Systems*, vol. 6, no. 1, pp. 19–36, 2006.

[16] N. Cai and R. W. Yeung, "Network error correction, part II: Lower bounds", *Commun. Information Systems*, vol. 6, no. 1, pp. 37–54, 2006.

[17] R. Matsumoto, "Construction algorithm for network error–correcting codes attaining the Singleton bound", *IEICE Trans. Fundamentals*, vol. E90–A, no. 9, pp. 1–7, Sept. 2007.

[18] S. Jaggi, M. Langberg, S. Katti, T. Ho, D. Katabi, M. Médard, and M. Effros, "Resilient network coding in the presence of byzantine adversaries", *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2596–2603, June 2008.

[19] H. Balli, X. Yan, and Z. Zhang, "On randomized linear network codes and their error correction capabilities", *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3148–3160, July 2009.

[20] T. Etzion and N. Silberstein, "Error–correcting codes in projective spaces via rank–metric codes and Ferrers diagrams", *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 2909–2919, July 2009.

[21] S.–T. Xia and F.–W. Fu, "Johnson type bounds on constant dimension codes", *Designs, Codes, and Cryptography*, vol. 50, no. 2, pp. 163–172, June 2008.

[22] E. M. Gabidulin and M. Bossert, "Codes for network coding", *IEEE Int. Symposium Inform. Theory*, pp. 867–870, July 2008.

[23] D. Silva, F. R. Kschischang, and R. Koetter, "A rank–metric approach to error control in random network coding", *IEEE Trans. Inform. Theory*, vol. 54, no. 9, pp. 3951–3967, Sept. 2008.

[24] A. Kohnert and S. Kurz, "Construction of large constant dimension codes with a prescribed minimum distance", *Lecture Notes in Computer Science, Springer*, pp. 31–42, Dec. 2008.

[25] R. Ahlswede and H. Aydinian, "On error control codes for random network coding", *IEEE Int. Workshop Network Coding Theory and Applications*, pp. 68–73, June 2009.

[26] M. Gadouleau and Z. Yan, "Bounds on covering codes with the rank metric", *arXiv.org*, June 2009. [Online]. Available: http://arxiv.org/PS_cache/arxiv/pdf/0809/0809.2968v2.pdf.

[27] D. Silva and F. R. Kschischang, "On metrics for error correction in network coding", *arXiv.org*, Aug. 2009. [Online]. Available: http://arxiv.org/PS_cache/arxiv/pdf/0805/0805.3824v4.pdf.

[28] A. Khaleghi and F. R. Kschischang, "Projective space codes for the injection metric", *arXiv.org*, Apr. 2009. [Online]. Available: http://arxiv.org/PS_cache/arxiv/pdf/0904/0904.0813v2.pdf.

[29] N. Chen, M. Gadouleau, and Z. Yan, "Rank metric decoder architectures for noncoherent error control in random network coding", *IEEE Workshop Sig. Process. Systems*, Oct. 2009.